ISSN: 2414-1895 DOI: 10.6919/ICJE.202604_12(4).0055

Hourly Heat Load Prediction for District Heating Systems based on Long Short-Term Memory Network

Huan Ding[1,*], Yiming Zhan[2]

[1] College of Science, School of Tianjin University of Commerce, Tianjin, 300134, China

[2] School of Statistics and Data Science, Nankai University, Tianjin, 300071, China

[*]Corresponding author: dingh069@139.com

Abstract

In conjunction with the rapid progress in intelligent heating technology, heat load-based prediction and on-demand heating have been considered as the most important factor for the realization of optimized regulation in a centralized system. The prediction of heat load in heating systems will be useful for the optimization of automatic control and green heating. But traditionally applied methods are less effective and vulnerable to noise, hence they cannot meet the demands of industrial practice. Recently, deep learning techniques have been intensively researched and used in order to extract meaningful features from the data and construct models. LSTM model has proven superior compared to other approaches in terms of performance. In this regard, the purpose of this study is to use a deep learning method of Long Short-Term Memory Network for the accurate heat load prediction in DHSs in order to reduce the waste of resources.

Keywords

Heat Load Prediction; DHSs; Deep Learning; LSTM Model.

1. Introduction

The energy consumption by buildings makes up a considerable share of the world's energy consumption, and the thermal energy consumption of buildings alone accounts for more than 40% of the entire societal energy consumption. According to statistics, the coverage area and coverage volume of centralized heating networks continue to expand in China. As reported, the total coverage distance of the centralized heating pipelines in Chinese cities was 461,000 km by 2021, covering an area of about 10.6 billion square meters [1]. Similarly, the energy consumption of heating in Chinese northern cities came up to 212 million tons of coal equivalent (tce) in 2021 [2]. Therefore, it is imperative to make accurate thermal load forecasting.

District heating systems on a large scale are important civil infrastructural projects in China and are marked by huge scales of the networks involved and high energy consumption. The main components of such a system include heat sources, network distribution, and consumers of the heat. Heat exchangers play an important role in such projects by serving as key points that connect the heat source to the consumer; hence, they affect the heat generated and the user's room temperature. Proper load forecasting for the control of such stations is necessary, and proper modeling will aid in achieving this.

Given the fast advancement of IoT and artificial intelligence techniques, the heating system of the future will most likely be geared towards being intelligent in nature, precise predictions, as well as automatic optimization and control. This is why the field of thermal load forecasting has received significant attention among energy studies professionals and computer scientists worldwide. Depending on the forecast objective, there are basically two types of thermal load forecasting. These include demand-side forecasting and supply-side forecasting [3]. Whereas the former deals with the thermal requirement on the side of the consumers and users, the latter involves the prediction of the thermal production capacity of heat sources and heat exchanges.

The evolution of thermal load forecasting methods can be classified into three stages. The first stage involved the use of statistical models, in which the scientists made use of their expertise in the field of thermal energy to develop statistical prediction models. Models like ARMA, SARMA, and MLR have been extensively used for time series forecasting because of their simplicity and ease of interpretation. But traditional statistical models are extremely sensitive to outliers and do not perform well in handling the multi-coupling behavior of thermal loads, making them unable to generalize properly.

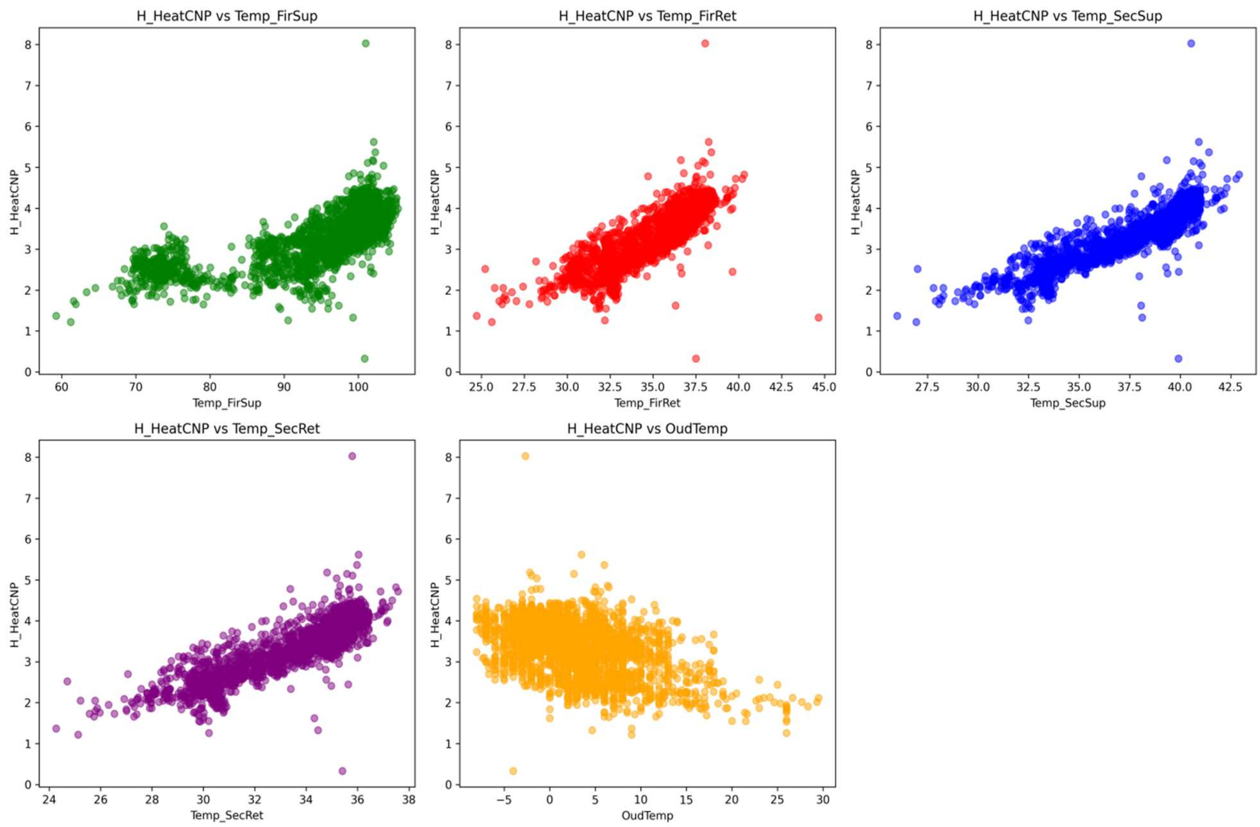

Fig. 1 Scatter plots showing hourly heat load and related variables

Stage two combines thermal energy information with computer techniques, mainly using traditional machine learning models. In stage three, which is mainly made possible by computer science development, modeling thermal load forecasting is conducted by using big data and artificial intelligence technologies. Application of Artificial Neural Network (ANN) has greatly improved accuracy in district heating load prediction leading to optimization in energy management[4]. Thermal load prediction plays a key part in energy efficiency improvement and demand regulation. In order to reduce energy consumption in residence heating, different approaches like time series prediction and regression analysis have been developed[5].

Xue et al. [6] proposed three machine learning algorithms (Support Vector Regression, Deep Neural Networks, and Extreme Gradient Boosting) for the thermal load prediction problem in the centralized heating system. Among them, XGBoost achieved the highest accuracy with an error margin of 9.59% in terms of MAPE. In another study, Yu et al. (2023) [7] compared five types of data-driven forecast models and concluded that the improved recurrent neural network, more specifically the long-short term memory (LSTM) model equipped with attention mechanism, was the most accurate in predicting

 Volume 12 Issue 4, 2026
DOI: 10.6919/ICJE.202604_12(4).0055

the heating and cooling load for the next 24 hours. As an artificial neural network architecture, the LSTM network has attracted considerable attention for its outstanding capability in capturing complicated nonlinear associations between data.

The remainder of this paper is organized as follows: Section 2 analyzes the characteristics of thermal load;Section 3 outlines the mathematical principles of the Long Short-Term Memory forecasting model; Section 4 presents the experimental setup and discusses the results; and Section 5concludes the study.

2. Feature Analysis

The DHS is a complicated system due to its high levels of non-linearity, variability in time, and time-lag effect. In this particular situation, the demand for heat by stations depends on various conditions such as the primary temperatures of supply and return water, secondary temperatures of supply and return water, pressure, and temperature of the outside air.The parameters of the influential factors are recorded in the historical database of the monitoring center with the help of temperature and pressure sensors. Hourly thermal load and the scatter plots of its influential factors at the heat exchanger station are provided in Fig. 1. As can be seen, there is evidence of non-linearity and multi-faceted correlations in thermal load. In this study, The Pearson Correlation Coefficient (PCC) is used to analyze the correlations of thermal load with each influential factor, as provided in Equation (1).

$$\rho = \frac{\sigma_{xy}}{\sigma_x \sigma_y} = \frac{\sum(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum(x_i - \bar{x})^2}\sqrt{\sum(y_i - \bar{y})^2}} \tag{1}$$

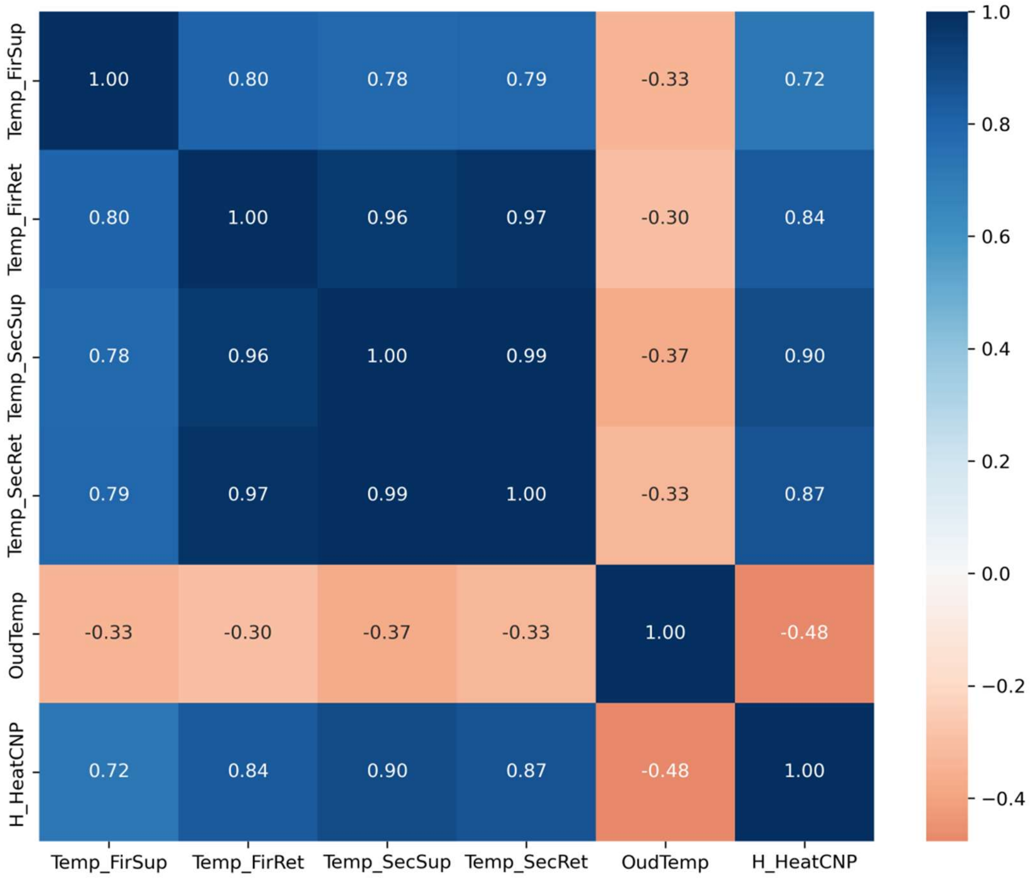

Fig. 2 Pearson correlation coefficient matrix chart between heat load and its affecting factors

ISSN: 2414-1895 DOI: 10.6919/ICJE.202604_12(4).0055

In the equation, σ_{xy} denotes the covariance of both variables X and Y within the population, σ_x represents the standard deviation of variable X within the population, while σ_y refers to the standard deviation of variable Y within the population. The findings of the PCC test between thermal load and its influencing parameters are displayed in Fig. 2, in which all PCC values are not less than 0.3. Hence, in this study, the supply temperature of primary water, return temperature of primary water, supply temperature of secondary water, return temperature of secondary water, and outdoor temperature are used as input data of the LSTM network.

3. Methodology and Analysis

3.1 Long Short-Term Memory Network

Long Short-Term Memory (LSTM) is an advanced form of Recurrent Neural Network used to solve the problem that existed with RNN with respect to the issue of long term dependency. With the use of “gates”, LSTM has been able to learn and remember information in long sequences of data, thus being capable of controlling the storage, update, and output of such information. Thus, LSTM has been successful in solving problems of vanishing gradients and exploding gradients. There are three gates in LSTM: the forget “gate”, “input gate” and “output gate”.

3.1.1 Forget Gate

Forget gate is responsible for determining which of the memory content needs to be held and which needs to be forgotten. Forget gate takes weighted summation of previous hidden state and input state and then applies the Sigmoid function to generate output values between [0,1], indicating the degree of forgetting.

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f). \tag{2}$$

Herein, f_t denotes the output from the forget gate, σ refers to the sigmoid activation function, h_{t-1} indicates the hidden state at the prior time step t, x_t shows the input in the current time stept, and W_f and b_f are the weight and bias, respectively.

3.1.2 Input Gate

The input gate will handle the updating of data input in the present time period. The sigmoid function will help decide the values that should be updated, while the tanh function will create candidate memory cells (data information).

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i). \tag{3}$$

$$\tilde{C}_t = \tanh \cdot (W_C \cdot [h_{t-1}, x_t] + b_C). \tag{4}$$

$$C_t = f_t \cdot C_{t-1} + i_t \cdot \tilde{C}_t. \tag{5}$$

Herein, i_t denotes the output from the input gate, $\tilde{C}_t$ is the memory candidate cell, and the C_t is updated under the control of the forget gate and the input gate.

3.1.3 Output Gate

The output gate will decide the output for that specific moment in time step t (hidden state ht). This will be achieved by evaluating the memory cell currently available to it (after applying Tanh activation) and the decision of the input gate.

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o). \tag{6}$$

ISSN: 2414-1895 DOI: 10.6919/ICJE.202604_12(4).0055

$$h_t = o_t \cdot \tanh(C_t). \tag{7}$$

3.2 Evaluation Criteria

Evaluation metrics such as Mean Squared Error (MSE), Root Mean Squared Error (RMSE), Mean Absolute Error (MAE), and Mean Absolute Percentage Error (MAPE) are chosen to gauge the accuracy of the forecasting of heat load. These formulas are expressed below:

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(\widehat{Q_i} - Q_i)^2. \tag{8}$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\widehat{Q_i} - Q_i)^2}. \tag{9}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|\widehat{Q_i} - Q_i\right|. \tag{10}$$

$$MAPE = \frac{100}{n}\sum_{i=1}^{n}\left|\frac{\widehat{Q_i} - Q_i}{Q_i}\right|. \tag{11}$$

Herein, $\widehat{Q_i}$ denotes the predicted thermal load at the ith time stamp, and Q_i denotes the observed thermal load at the ith time stamp.

4. Experimental Results and Discussion

4.1 Data Segmentation Rules

This paper adopts two different datasets; namely, the training dataset and the test dataset. As heat loads belong to the type of time series data, there are obvious logical relations between adjacent data points within the series. Hence, random division of the dataset using the randomly assigned number seed should not be employed, because it would violate the logical relation within the training dataset. The last 24 points of data are considered the testing dataset while the rest of the data points will be considered the training dataset to forecast the future 24 hours of heat load predictions. Moreover, it is critical that the prediction model keeps getting updated in order to continue being accurate. The update frequency of the LSTM Model prediction model proposed in this paper is set to be one update per day, meaning that the prediction model is trained at midnight each day. The updated training dataset contains all aggregated data of hourly heat load records since the beginning of the heating season up to now.

4.2 Prediction Results and Performance Comparison

For the evaluation of prediction performances of LSTM model, the following study includes the comparative analysis between the proposed method and some other sophisticated methods. Also, for this comparative experiment, historical hour data of heat exchange station are used in this study. The results of thermal load predictions and comparative experiments are given in fig. 3. In order to perform a comprehensive evaluation for the performance of LSTM model, four different advanced algorithms named as RNN, TCN, AdaBoost and SVR are chosen in this study.

 Volume 12 Issue 4, 2026
DOI: 10.6919/ICJE.202604_12(4).0055

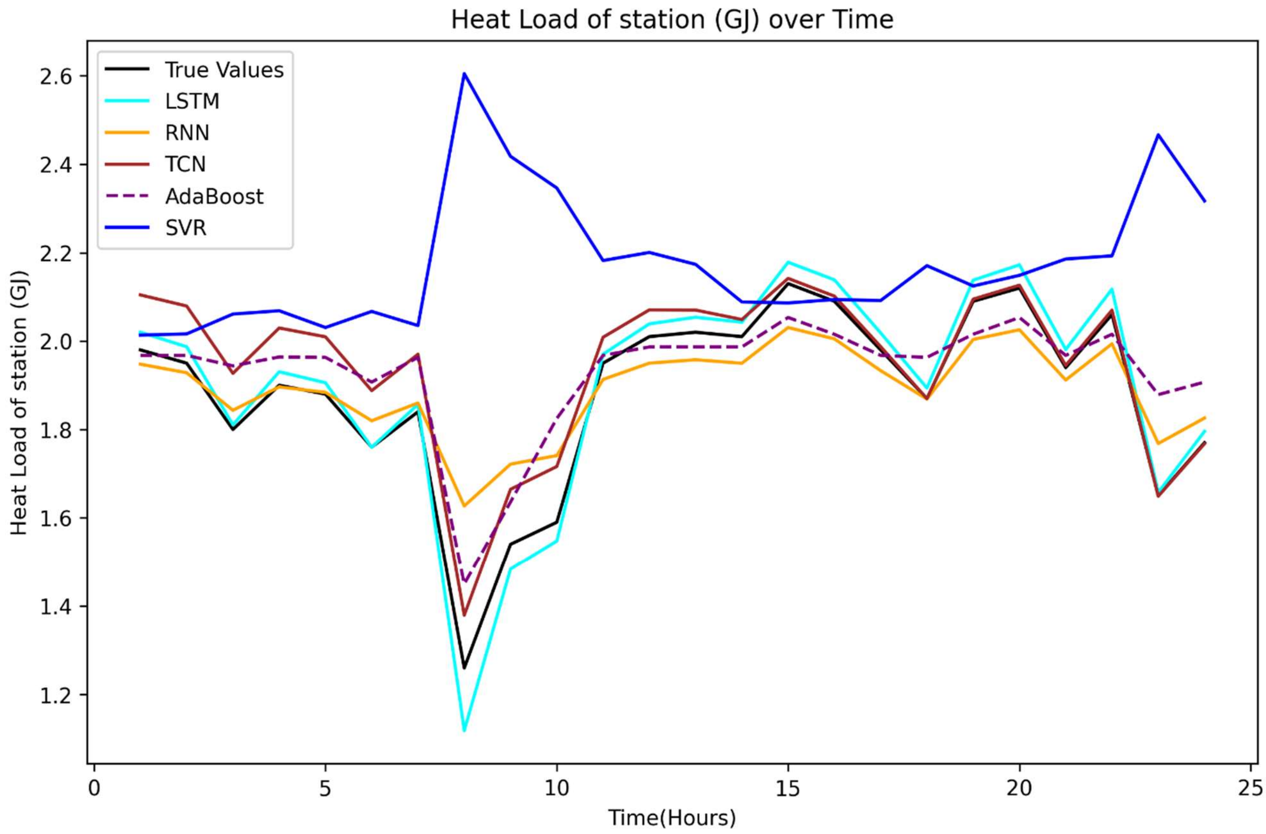

Fig. 3 Comparison chart of hourly heat loads actual values vs. predicted values of different algorithms in the test set (Unit: GJ)

The advantage of the LSTM model in prediction accuracy is apparent in the curve chart for the results of heat load prediction, as illustrated in Figure below. The prediction curve generated by the SVM algorithm is observed to fluctuate largely from the real values. The prediction curves produced by algorithms such as RNN, TCN, and AdaBoost also have some fluctuations from one another and have a lower degree of fitting than the prediction curve of the LSTM model. This implies that the LSTM model is capable of accurately identifying the features and variations of the thermal load through data-driven technology.

Table 1. Comparison of various algorithms'performance parameters statistically

	LSTM	RNN	TCN	AdaBoost	SVR
RMSE	0.046	0.106	0.085	0.109	0.447
MAE	0.038	0.074	0.064	0.088	0.306
MAPE	0.021	0.044	0.036	0.051	0.187
MSE	0.002	0.011	0.007	0.012	0.200

The statistic data of the prediction of heat load per hour through various methods is shown in Table 1. For the LSTM model, the statistical values of the heat exchange station including RMSE, MSE, MAE, and MAPE are 0.046, 0.002, 0.038, and 0.021. In comparison to RNN, TCN, AdaBoost, and SVR models, the accuracy of prediction of the LSTM model is enhanced by 56.6%, 45.9%, 57.8%, and 89.8% respectively. As can be seen from the experiment results, the prediction model based on LSTM model used in this paper has the strong feature extraction capability.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202604_12(4).0055

5. Summary

An accurate heat load prediction model is the indispensable technological basis that makes it possible for the SDHS to accomplish its ultimate goals. From the perspective of energy management, the heat load prediction model is capable of not only promoting intelligent innovation, but also offering guidance to heating facility design, ensuring the sustainable development of heating systems, and ensure the economical and efficient operation of heating systems in different aspects. The proposed LSTM heat load prediction model on the heat exchange station performs excellently, while the data required for the input variables can be easily acquired, and thus is quite promotive. In future studies, indoor comfort is expected to become a new input variable in the model for better practical application of the model.

References

[1] National Bureau of Statistics of China. China statistics year book [M]. Beijing: China Statistics Press; 2022.

[2] Building energy conservation research center, Tsinghua university. China building energy efficiency annual development research report 2023 (urban energy system topic) [M]. Beijing: China Architecture & Building Press; 2023.

[3] Q. Zhang, Z. Tian, Z. Ma, G. Li, Y. Lu., J. Niu, Development of the heating load prediction model for the residential building of district heating based on model calibration, Energy 205(2020) 117949.

[4] Novitsky NN, Shalaginova ZI, et al. Smarter smart district heating. Proc IEEE 2020;108(9):1596–611.

[5] Zhang J, Hu Y, Yuan Y, et al. Accuracy improvement of the load forecasting in the district heating system by the informer-based framework with the optimal step size selection. Energy 2024;291:130347.

[6] Puning Xue, Yi Jiang, Zhou Zhigang, et al. Multi-step ahead forecasting of heat load in district heating systems using machine learning algorithms. Energy 2019;188: 116085.

[7] YU H, ZHONG F, DU Y, et al. Short-term cooling and heating loads forecasting of building district energy system based on data-driven models [J]. Energy and Buildings, 2023, 298(000): 15.